EXHIBIT 21

Page 1

IN THE UNITED STATES DISTRICT COURT

FOR THE NORTHERN DISTRICT OF CALIFORNIA

SAN FRANCISCO DIVISION

_____

IN RE MATTER OF:)

RICHARD KADREY, et al.,   )

Plaintiff,   )

vs.   ) C.A. NO.:

META PLATFORMS, INC.,   ) 3:23-cv-03417-VC

Defendant.   )

_____)

VIDEOTAPED DEPOSITION OF YANN LeCUN, Ph.D.

Palo Alto, California

Thursday, November 21, 2024

** HIGHLY CONFIDENTIAL - ATTORNEYS EYES ONLY **

UNDER THE PROTECTIVE ORDER

Stenographically Reported by:

HEATHER J. BAUTISTA, CSR, CRR, RPR, CLR

Realtime Systems Administrator

California CSR License #11600

Oregon CSR License #21-0005

Washington License #21009491

Nevada CCR License #980

Texas CSR License #10725

_____

DIGITAL EVIDENCE GROUP

1730 M. Street, NW, Suite 812

Washington, D.C. 20036

(202) 232-0646

Page 49

```
 1              THE WITNESS:  I don't know if it's true.

 2      Q.   (By Ms. Geman)  You don't know if it's true

 3   that there are Meta documents that say that or you

 4   don't know if it's true that the book datasets, in

 5   fact, are among the most valuable?

 6              MR. WEINSTEIN:  Same objection.

 7              THE WITNESS:  The latter.

 8      Q.   (By Ms. Geman)  Okay.

 9              So you do know that some -- Meta has

10   said -- Meta personnel have made that assertion;

11   correct?

12      A.   No.

13      Q.   You're saying they haven't made it or

14   you're just not aware one way or the other?

15              MR. WEINSTEIN:  Object to form.

16              THE WITNESS:  I'm not aware.

17      Q.   (By Ms. Geman)  And why is training on long

18   text important?

19      A.   So large language models have -- it's

20   called a context window which is the number of words

21   or token technically that it can use to predict the

22   next words in the text, and current models like
```

Page 50

```
 1    LLaMA 3 have a context window of 128,000 tokens

 2    which is on the order of a hundred thousand words,

 3    roughly.

 4           And so to take full advantage of the

 5    redundancy in the text, you have to have texts that

 6    are at least that long.  They can be from multiple

 7    authors on the same topic.

 8      Q.   And why is context window important?

 9      A.   Well, the context window is -- in current

10    LLMs, the context window is basically the only

11    memory the system has; working memory, I would say.

12    It has memory in the weights, but the sort of

13    working memory is the context.

14      Q.   And LLaMA cannot go beyond its pre-training

15    context window; is that correct?

16      A.   It's -- it's hard-wired in the architecture

17    so it's 128,000 --

18           (Stenographer clarification.)

19           THE WITNESS:  -- for LLaMA 3 and I believe

20    for LLaMA 1 it was 32,000, if I remember correctly.

21      Q.   (By Ms. Geman)  Is the desire to keep

22    increasing the context window?
```

Page 185

```
 1   blueprint mean training on sort of massive amounts

 2   of data to -- or sort of training on bigger and

 3   bigger sets of data without kind of looking more at

 4   the relationship between the nodes?

 5        A.    No.   The blueprint refers to the

 6   autoregressive LLM architecture.

 7        Q.    Okay.

 8              Which aspect of the architecture?

 9        A.    Well, the fact that the way an LLM is

10   trained is, you -- you build a neural network, a

11   deep-learning system, particularly -- particularly

12   architecture called a transformer, but that's

13   actually irrelevant to the point.

14              You -- you give it a piece of text and then

15   you train it to just reproduce the text on its

16   output, but it can only look at words that are to

17   the left of a particular word it's trying to

18   predict.  And so in the end, what the system can do,

19   is given a segment of text, it can produce a guess

20   as to what the next word in that text is going to

21   be.  In fact, it cannot exactly predict what word

22   comes next so it's going to produce a --
```

Page 323

1   I, HEATHER J. BAUTISTA, CSR No. 11600,

2   Certified Shorthand Reporter, certify:

3   That the foregoing proceedings were taken

     before me at the time and place therein set forth,

4   at which time the witness declared under penalty of

     perjury; that the testimony of the witness and all

5   objections made at the time of the examination were

     recorded stenographically by me and were thereafter

6   transcribed under my direction and supervision; that

     the foregoing is a full, true, and correct

7   transcript of my shorthand notes so taken and of the

8   testimony so given;

9   (XX) Reading and signing was not requested/offered.

10   I further certify that I am not financially

11   interested in the action, and I am not a relative or

12   employee of any attorney of the parties, nor of any

13   of the parties.

14   I declare under penalty of perjury under the

15   laws of California that the foregoing is true and

16   correct.   Dated: November 26, 2024

17

18

19

20

21

22
                    HEATHER J. BAUTISTA, CSR, CRR, RPR, CLR